WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report (i)

Jun 14, 2020 – 03:28 pm BST

PDB ID	:	2N18
Title	:	Dominant form of the low-affinity complex of yeast cytochrome c and cytochrome c peroxidase
Authors	:	Volkov, A.; Van de Water, K.
Deposited on	:	2015-03-24

The following versions of software and data (see references (i)) were used in the production of this report:

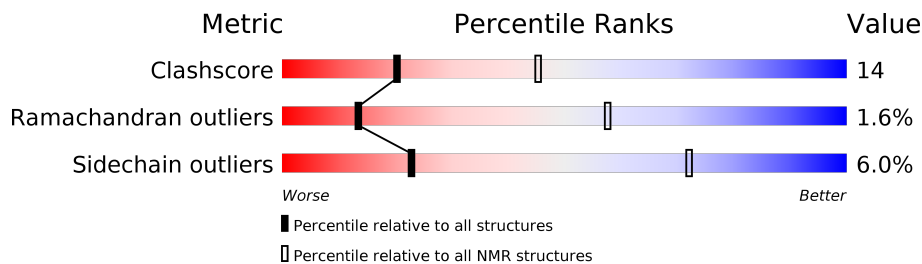|		
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 7%.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Clashscore		14
Ramachandran outliers		1.6%
Sidechain outliers		6.0%

Worse   *Better*

■ Percentile relative to all structures
▯ Percentile relative to all NMR structures

Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	294	89%　　　　9%　•
2	B	108	73%　　　　24%　•
3	C	108	46%　　　39%　　14%　•

2 Ensemble composition and analysis (i)

This entry contains 15 models. Model 2 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:1-A:294, B:-5-B:-1, B:1-B:103 (402)	-0.00	2
2	C:-5-C:-1, C:1-C:103 (108)	0.00	2

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

NmrClust was unable to cluster the ensemble.

Error message: 1 Unexpected character

3 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 8241 atoms, of which 4046 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Cytochrome c peroxidase, mitochondrial.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	294	4621	1512	2249	395	459	6	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	128	ALA	CYS	ENGINEERED MUTATION	UNP P00431
A	197	CYS	VAL	ENGINEERED MUTATION	UNP P00431

- Molecule 2 is a protein called Cytochrome c iso-1.

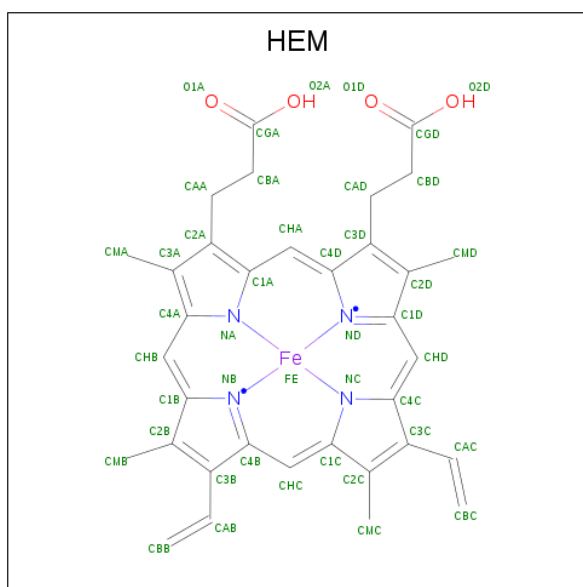| Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
2	B	108	1698	534	851	151	157	5	0

There are 3 discrepancies between the modelled and reference sequences:

| Chain	Residue	Modelled	Actual	Comment	Reference
B	-5	ALA	-	EXPRESSION TAG	UNP P00044
B	81	CYS	ALA	ENGINEERED MUTATION	UNP P00044
B	102	THR	CYS	ENGINEERED MUTATION	UNP P00044

- Molecule 3 is a protein called Cytochrome c iso-1.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
3	C	108	1699	534	852	151	157	5	0

- Molecule 4 is PROTOPORPHYRIN IX CONTAINING FE (three-letter code: HEM) (formula: $C_{34}H_{32}FeN_4O_4$).

HEM

Mol	Chain	Residues	Atoms					
			Total	C	Fe	H	N	O
4	A	1	73	34	1	30	4	4

- Molecule 5 is HEME C (three-letter code: HEC) (formula: $C_{34}H_{34}FeN_4O_4$).

HEC
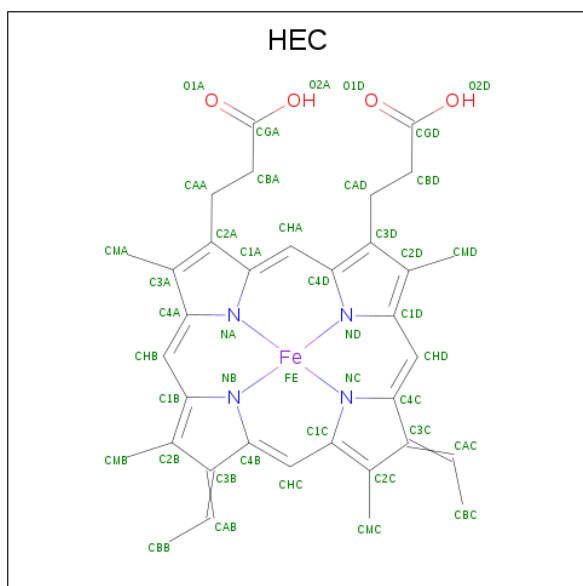
| Mol	Chain	Residues	Atoms					
			Total	C	Fe	H	N	O
5	B	1	75	34	1	32	4	4
5	C	1	Total	C	Fe	H	N	O
			75	34	1	32	4	4

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

- Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A:

- Molecule 2: Cytochrome c iso-1

Chain B:

- Molecule 3: Cytochrome c iso-1

Chain C:
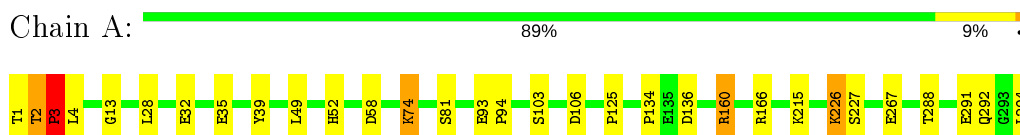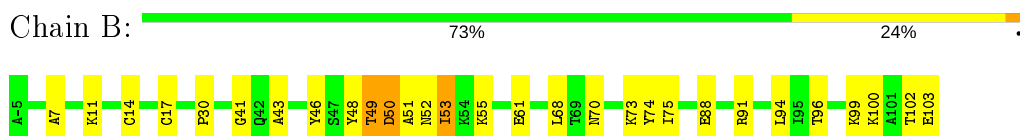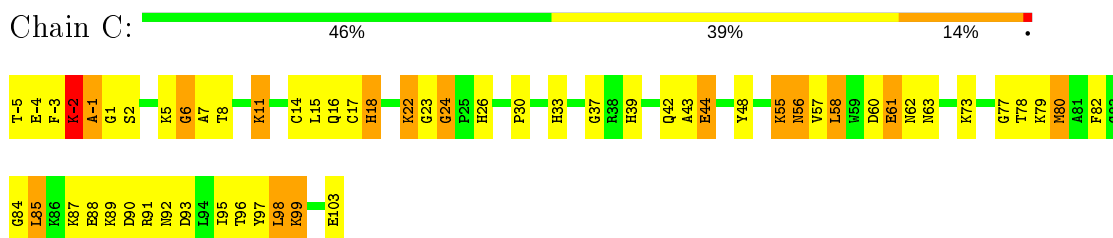
4.2 Scores per residue for each member of the ensemble

Colouring as in section 4.1 above.

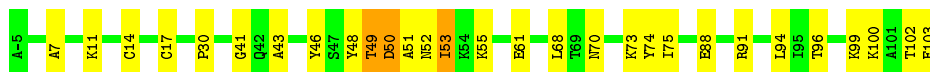### 4.2.1 Score per residue for model 1
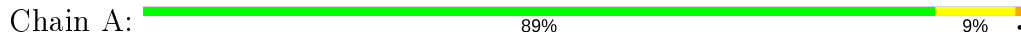
- Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A:

T1 T2 P3 L4 G13 L28 E32 E35 Y39 L49 H52 D58 K74 S81 E93 P94 S103 D106 P125 P134 E135 D136 R160 R166 K215 K226 S227 E267 T288 E291 Q292 G293 L294

• Molecule 2: Cytochrome c iso-1

Chain B: 73% 24%

A-5 A7 K11 C14 C17 P30 G41 Q42 A43 Y46 S47 Y48 T49 D50 A51 N52 I53 K54 K55 E61 L68 T69 N70 K73 Y74 I75 E88 R91 L94 I95 T96 K99 K100 A101 T102 E103

• Molecule 3: Cytochrome c iso-1

Chain C: 46% 39% 14%

T-5 E-4 F-3 K-2 A-1 G1 S2 K5 G6 A7 T8 K11 C14 L15 Q16 C17 H18 K22 G23 G24 P25 H26 P30 H33 G37 R38 H39 Q42 A43 E44 Y48 K55 N56 V57 L58 W59 D60 E61 N62 N63 K73 G77 T78 K79 M80 A81 F82 G83

G84 L85 K86 K87 E88 K89 D90 R91 N92 D93 L94 I95 T96 Y97 L98 K99 E103

4.2.2 Score per residue for model 2 (medoid)

• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A: 89% 9%

T1 T2 P3 L4 G13 L28 E32 E35 Y39 L49 H52 D58 K74 S81 E93 P94 S103 D106 P125 P134 E135 D136 R160 R166 K215 K226 S227 E267 T288 E291 Q292 G293 L294

• Molecule 2: Cytochrome c iso-1

Chain B: 73% 24%

A-5 A7 K11 C14 C17 P30 G41 Q42 A43 Y46 S47 Y48 T49 D50 A51 N52 I53 K54 K55 E61 L68 T69 N70 K73 Y74 I75 E88 R91 L94 I95 T96 K99 K100 A101 T102 E103

• Molecule 3: Cytochrome c iso-1

Chain C: 46% 39% 14%

T-5 E-4 F-3 K-2 A-1 G1 S2 K5 G6 A7 T8 K11 C14 L15 Q16 C17 H18 K22 G23 G24 P25 H26 P30 H33 G37 R38 H39 Q42 A43 E44 Y48 K55 N56 V57 L58 W59 D60 E61 N62 N63 K73 G77 T78 K79 M80 A81 F82 G83

G84 L85 K86 K87 E88 K89 D90 R91 N92 D93 L94 I95 T96 Y97 L98 K99 E103

4.2.3 Score per residue for model 3

• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A: 89% 9%

T1 T2 P3 L4 G13 L28 E32 E35 Y39 L49 H52 D58 K74 S81 E93 P94 S103 D106 P125 P134 E135 D136 R160 R166 K215 K226 S227 E267 T288 E291 Q292 G293 L294

• Molecule 2: Cytochrome c iso-1

Chain B: 73% 24%

A-5 A7 K11 C14 C17 P30 G41 Q42 A43 Y46 S47 Y48 T49 D50 A51 N52 I53 K54 K55 E61 L68 T69 N70 K73 Y74 I75 E88 R91 L94 I95 T96 K99 K100 A101 T102 E103

• Molecule 3: Cytochrome c iso-1

Chain C: 46% 39% 14%

T-5 E-4 F-3 K-2 A-1 G1 S2 K5 G6 A7 T8 K11 C14 L15 Q16 C17 H18 K22 G23 G24 P25 H26 P30 H33 G37 R38 H39 Q42 A43 E44 Y48 K55 N56 V57 L58 W59 D60 E61 N62 N63 K73 G77 T78 K79 M80 A81 F82 G83

G84 L85 K86 K87 E88 K89 D90 R91 N92 D93 L94 I95 T96 Y97 L98 K99 E103

4.2.4 Score per residue for model 4

• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A: 89% 9%

T1 T2 P3 L4 G13 L28 E32 E35 Y39 L49 H52 D58 K74 S81 E93 P94 S103 D106 P125 P134 E135 D136 R160 R166 K215 K226 S227 E267 T288 E291 Q292 G293 L294

• Molecule 2: Cytochrome c iso-1

Chain B: 73% 24%

A-5 A7 K11 C14 C17 P30 G41 Q42 A43 Y46 S47 Y48 T49 D50 A51 N52 I53 K54 K55 E61 L68 T69 N70 K73 Y74 I75 E88 R91 L94 I95 T96 K99 K100 A101 T102 E103

• Molecule 3: Cytochrome c iso-1

Chain C: 46% 39% 14%

T-5 E-4 F-3 K-2 A-1 G1 S2 K5 G6 A7 T8 K11 C14 L15 Q16 C17 H18 K22 G23 G24 P25 H26 P30 H33 G37 R38 H39 Q42 A43 E44 Y48 K55 N56 V57 L58 W59 D60 E61 N62 N63 K73 G77 T78 K79 M80 A81 F82 G83

G84 L85 K86 K87 E88 K89 D90 R91 N92 D93 L94 I95 T96 Y97 L98 K99 E103

4.2.5 Score per residue for model 5

• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A:

89% 9%

• Molecule 2: Cytochrome c iso-1

Chain B:

73% 24%

• Molecule 3: Cytochrome c iso-1

Chain C:

46% 39% 14%
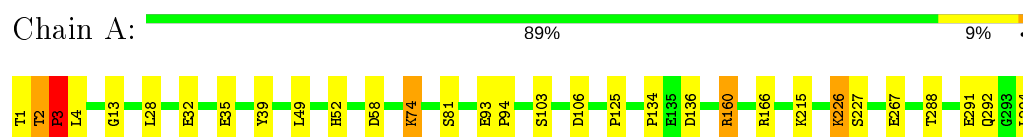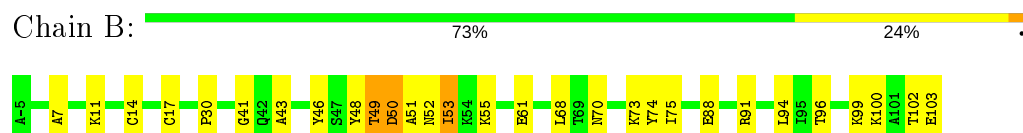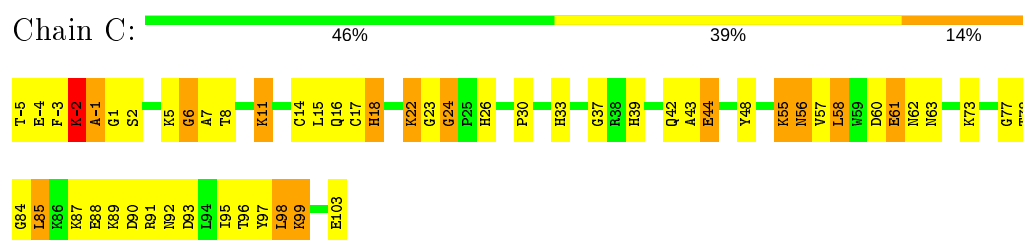
4.2.6 Score per residue for model 6

• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A:

89% 9%

• Molecule 2: Cytochrome c iso-1

Chain B:

73% 24%

• Molecule 3: Cytochrome c iso-1

Chain C:

46% 39% 14%

G84 L85 K86 K87 E88 K89 D90 R91 N92 D93 L94 I95 T96 Y97 L98 K99 E103
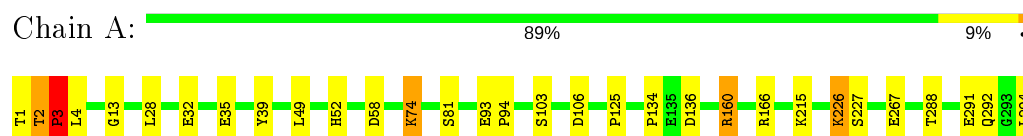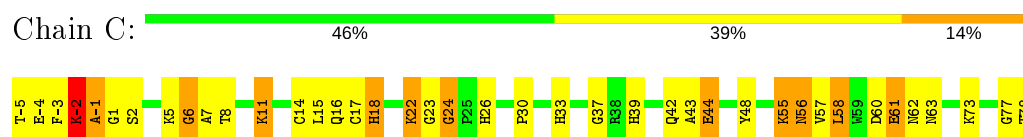
4.2.7 Score per residue for model 7

• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A: 89% 9%

T1 T2 P3 L4 G13 L28 E32 E35 Y39 L49 H52 D58 K74 S81 E93 P94 S103 D106 P125 P134 E135 D136 R160 R166 K215 K226 S227 E267 T288 E291 Q292 G293 L294

• Molecule 2: Cytochrome c iso-1

Chain B: 73% 24%

A-5 A7 K11 C14 C17 P30 G41 Q42 A43 Y46 S47 Y48 T49 D50 A51 N52 I53 K54 K55 E61 L68 T69 N70 K73 Y74 I75 E88 R91 L94 I95 T96 K99 K100 A101 T102 E103

• Molecule 3: Cytochrome c iso-1

Chain C: 46% 39% 14%

T-5 E-4 F-3 K-2 A-1 G1 S2 K5 G6 A7 T8 K11 C14 L15 Q16 C17 H18 K22 G23 G24 P25 H26 P30 H33 G37 R38 H39 Q42 A43 E44 Y48 K55 N56 V57 L58 W59 D60 E61 N62 N63 K73 G77 T78 K79 M80 A81 F82 G83

G84 L85 K86 K87 E88 K89 D90 R91 N92 D93 L94 I95 T96 Y97 L98 K99 E103

4.2.8 Score per residue for model 8

• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A: 89% 9%

T1 T2 P3 L4 G13 L28 E32 E35 Y39 L49 H52 D58 K74 S81 E93 P94 S103 D106 P125 P134 E135 D136 R160 R166 K215 K226 S227 E267 T288 E291 Q292 G293 L294

• Molecule 2: Cytochrome c iso-1

Chain B: 73% 24%

A-5 A7 K11 C14 C17 P30 G41 Q42 A43 Y46 S47 Y48 T49 D50 A51 N52 I53 K54 K55 E61 L68 T69 N70 K73 Y74 I75 E88 R91 L94 I95 T96 K99 K100 A101 T102 E103

• Molecule 3: Cytochrome c iso-1

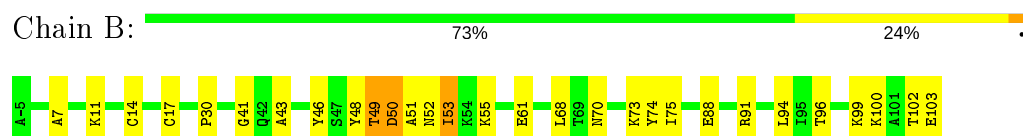Chain C: 47% 38% 14%
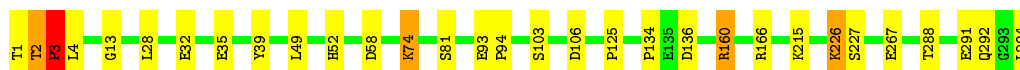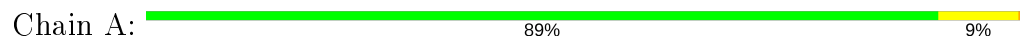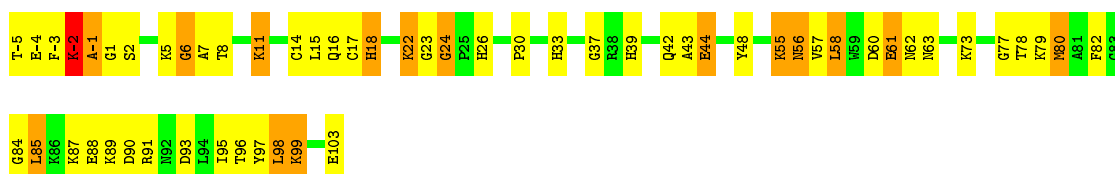
4.2.9 Score per residue for model 9

• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A:
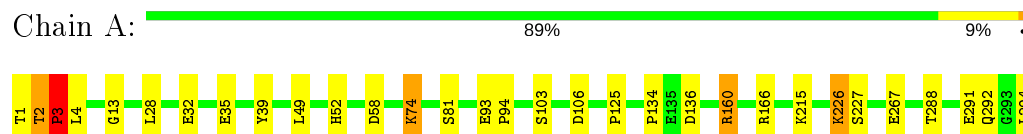
• Molecule 2: Cytochrome c iso-1

Chain B:

• Molecule 3: Cytochrome c iso-1
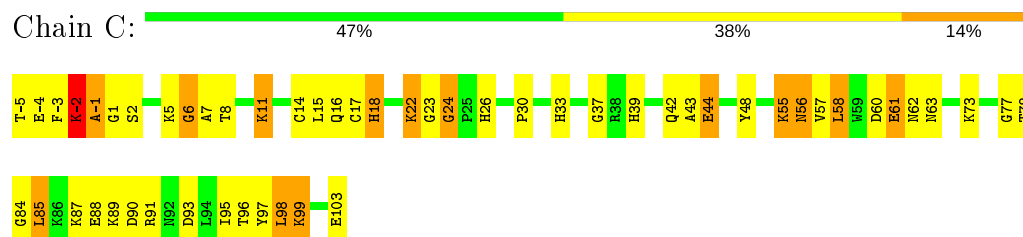
Chain C:

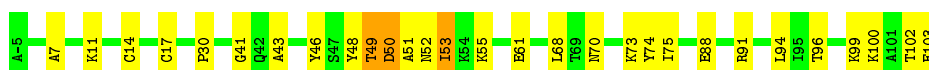
4.2.10 Score per residue for model 10

• Molecule 1: Cytochrome c peroxidase, mitochondrial
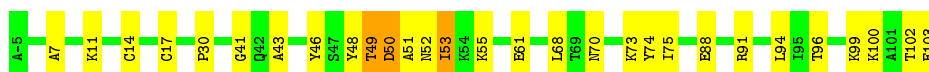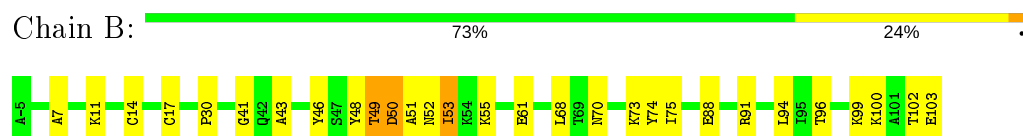
Chain A:

• Molecule 2: Cytochrome c iso-1

Chain B:

- Molecule 3: Cytochrome c iso-1

Chain C: 51% 35% 13%

T-5 E-4 F-3 K-2 A-1 G1 S2 K5 G6 A7 T8 K11 C14 L15 Q16 C17 H18 K22 G23 G24 P30 H33 G37 R38 H39 Q42 A43 E44 Y48 K55 N56 V57 L58 E61 N62 N63 K73 K79 M80 A81 F82 G83 G84 L85 K86 K87 E88

K89 D90 R91 N92 D93 L94 I95 T96 Y97 L98 K99 E103

4.2.11   Score per residue for model 11

- Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A: 89% 9%

T1 T2 P3 L4 G13 L28 E32 E35 Y39 L49 H52 D58 K74 S81 E93 P94 S103 D106 P125 P134 E135 D136 R160 R166 K215 K226 S227 E267 T288 E291 Q292 G293 L294

- Molecule 2: Cytochrome c iso-1

Chain B: 73% 24%

A-5 A7 K11 C14 C17 P30 G41 Q42 A43 Y46 S47 Y48 T49 D50 A51 N52 I53 K54 K55 E61 L68 T69 N70 K73 Y74 I75 E88 R91 L94 I95 T96 K99 K100 A101 T102 E103

- Molecule 3: Cytochrome c iso-1

Chain C: 49% 37% 13%

T-5 E-4 F-3 K-2 A-1 G1 S2 K5 G6 A7 T8 K11 C14 L15 Q16 C17 H18 K22 G23 G24 P30 H33 G37 R38 H39 Q42 A43 E44 Y48 K55 N56 V57 L58 W59 D60 E61 N62 N63 K73 K79 M80 A81 F82 G83 G84 L85 K86 K87

E88 K89 D90 R91 N92 D93 L94 I95 T96 Y97 L98 K99 E103

4.2.12   Score per residue for model 12

- Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A: 89% 9%

T1 T2 P3 L4 G13 L28 E32 E35 Y39 L49 H52 D58 K74 S81 E93 P94 S103 D106 P125 P134 E135 D136 R160 R166 K215 K226 S227 E267 T288 E291 Q292 G293 L294

- Molecule 2: Cytochrome c iso-1

Chain B: 73% 24%

A-5 A7 K11 C14 C17 P30 G41 Q42 A43 Y46 S47 Y48 T49 D50 A51 N52 I53 K54 K55 E61 L68 T69 N70 K73 Y74 I75 E88 R91 L94 I95 T96 K99 K100 A101 T102 E103

• Molecule 3: Cytochrome c iso-1

Chain C: 49% 37% 13% •

T-5 E-4 F-3 K-2 A-1 G1 S2 K5 G6 A7 T8 K11 C14 L15 Q16 C17 H18 K22 G23 G24 P30 H33 G37 R38 H39 Q42 A43 E44 Y48 K55 N56 V57 L58 W59 D60 E61 N62 N63 K73 K79 M80 A81 F82 G83 G84 L85 K86 K87

E88 K89 D90 R91 N92 D93 L94 I95 T96 Y97 L98 K99 E103

4.2.13 Score per residue for model 13

• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A: 89% 9% •

T1 T2 P3 L4 G13 L28 E32 E35 Y39 L49 H52 D58 K74 S81 E93 P94 S103 D106 P125 P134 E135 D136 R160 R166 K215 K226 S227 E267 T288 E291 Q292 G293 L294

• Molecule 2: Cytochrome c iso-1

Chain B: 73% 24% •

A-5 A7 K11 C14 C17 P30 G41 Q42 A43 Y46 S47 Y48 T49 D50 A51 N52 I53 K54 K55 E61 L68 T69 N70 K73 Y74 I75 E88 R91 L94 I95 T96 K99 K100 A101 T102 E103

• Molecule 3: Cytochrome c iso-1

Chain C: 46% 39% 14% •

T-5 E-4 F-3 K-2 A-1 G1 S2 K5 G6 A7 T8 K11 C14 L15 Q16 C17 H18 K22 G23 G24 P25 H26 P30 H33 G37 R38 H39 Q42 A43 E44 Y48 K55 N56 V57 L58 W59 D60 E61 N62 N63 K73 G77 T78 K79 M80 A81 F82 G83

G84 L85 K86 K87 E88 K89 D90 R91 N92 D93 L94 I95 T96 Y97 L98 K99 E103

4.2.14 Score per residue for model 14

• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A: 89% 9% •

T1 T2 P3 L4 G13 L28 E32 E35 Y39 L49 H52 D58 K74 S81 E93 P94 S103 D106 P125 P134 E135 D136 R160 R166 K215 K226 S227 E267 T288 E291 Q292 G293 L294

• Molecule 2: Cytochrome c iso-1

Chain B:

• Molecule 3: Cytochrome c iso-1

Chain C:
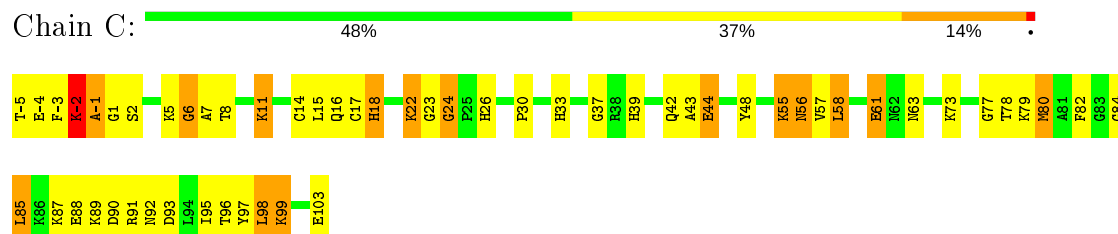
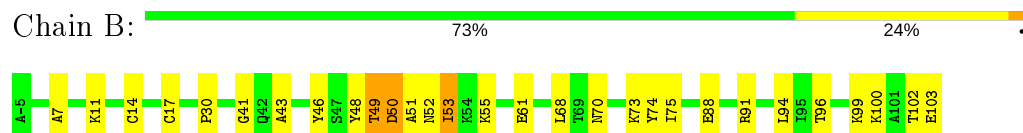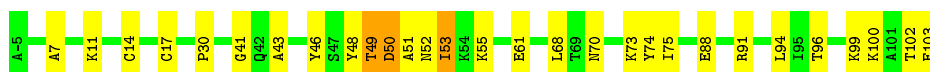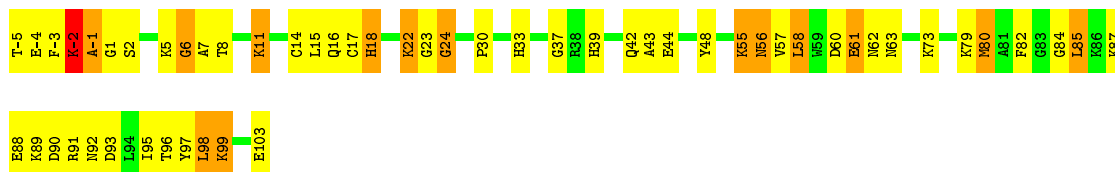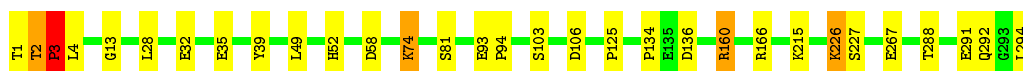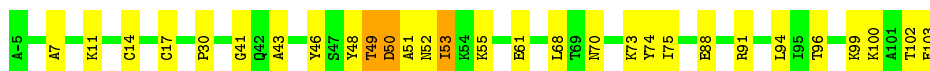### 4.2.15   Score per residue for model 15

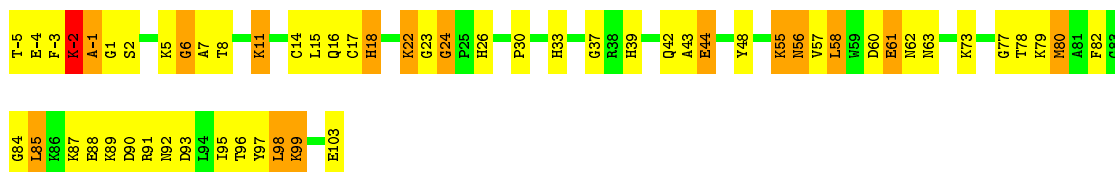• Molecule 1: Cytochrome c peroxidase, mitochondrial

Chain A:

• Molecule 2: Cytochrome c iso-1
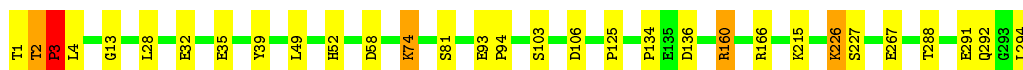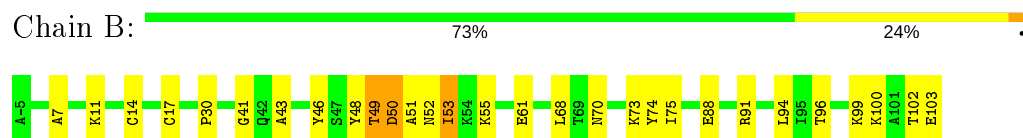
Chain B:

• Molecule 3: Cytochrome c iso-1

Chain C:
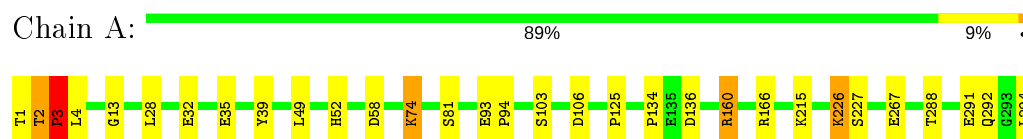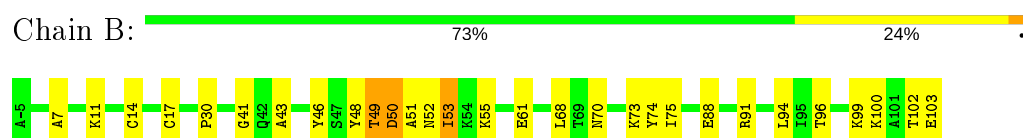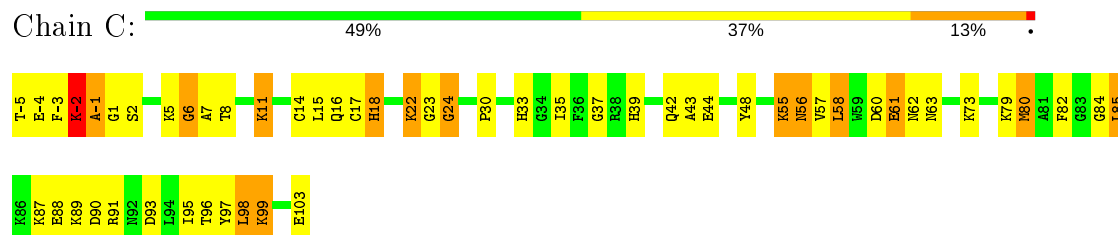
5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing*.

Of the 100 calculated structures, 15 were deposited, based on the following criterion: *structures with the lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
X-PLOR NIH	structure solution	
X-PLOR NIH	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	input_cs.cif
Number of chemical shift lists	1
Total number of shifts	458
Number of shifts mapped to atoms	458
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	7%

No validations of the models with respect to experimental NMR restraints is performed at this time.

6 Model quality (i)

6.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: HEM, HEC

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.60±0.00	0±0/2439 (0.0± 0.0%)	0.75±0.00	2±0/3304 (0.1± 0.0%)
2	B	0.52±0.00	0±0/865 (0.0± 0.0%)	0.73±0.00	0±0/1156 (0.0± 0.0%)
3	C	0.93±0.00	0±0/865 (0.0± 0.0%)	1.71±0.00	4±0/1156 (0.3± 0.0%)
All	All	0.67	0/62535 (0.0%)	1.02	87/84240 (0.1%)

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
3	C	85	LEU	CA-CB-CG	7.84	133.34	115.30	11	15
3	C	18	HIS	CA-CB-CG	7.75	126.78	113.60	14	15
1	A	160	ARG	NE-CZ-NH2	-6.82	116.89	120.30	1	15
3	C	44	GLU	CA-CB-CG	5.92	126.43	113.40	4	15
1	A	160	ARG	NE-CZ-NH1	5.56	123.08	120.30	1	15
3	C	92	ASN	CB-CA-C	5.03	120.46	110.40	13	12

There are no chirality outliers.

There are no planarity outliers.

6.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	2372	2249	2244	20±0
2	B	847	851	850	32±0
3	C	847	852	851	60±1
5	B	43	32	32	10±0
5	C	43	32	32	12±0
All	All	62925	60690	60585	1678

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 14.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:14:CYS:SG	5:B:201:HEC:CAB	1.29	2.21	14	15
2:B:17:CYS:SG	5:B:201:HEC:CAC	1.21	2.28	14	15
3:C:14:CYS:SG	5:C:201:HEC:CAB	1.21	2.29	8	15
3:C:17:CYS:SG	5:C:201:HEC:CAC	1.11	2.39	14	15
3:C:8:THR:HA	3:C:11:LYS:HD2	0.99	1.34	8	15
3:C:87:LYS:HB2	3:C:90:ASP:OD2	0.96	1.61	15	15
3:C:8:THR:HA	3:C:11:LYS:CD	0.95	1.91	1	15
3:C:14:CYS:SG	5:C:201:HEC:CBB	0.91	2.59	15	15
2:B:14:CYS:SG	5:B:201:HEC:HAB	0.81	2.14	14	15
2:B:14:CYS:SG	5:B:201:HEC:CBB	0.80	2.68	14	15
3:C:22:LYS:HD2	3:C:33:HIS:CD2	0.80	2.12	10	15
2:B:17:CYS:SG	5:B:201:HEC:C3C	0.80	2.69	14	15
3:C:17:CYS:SG	5:C:201:HEC:C3C	0.79	2.71	15	15
2:B:55:LYS:HG3	2:B:75:ILE:HG12	0.78	1.54	14	15
2:B:17:CYS:SG	5:B:201:HEC:CBC	0.73	2.76	14	15
3:C:17:CYS:SG	5:C:201:HEC:CBC	0.72	2.78	8	15
1:A:58:ASP:HA	1:A:160:ARG:HD3	0.72	1.62	14	15
2:B:49:THR:CG2	2:B:52:ASN:H	0.70	2.00	14	15
2:B:50:ASP:N	2:B:50:ASP:OD1	0.68	2.23	14	9
2:B:50:ASP:OD1	2:B:50:ASP:N	0.68	2.23	10	6
2:B:55:LYS:HE3	2:B:74:TYR:O	0.68	1.87	14	15
3:C:80:MET:HB2	5:C:201:HEC:C1D	0.67	2.20	7	15
3:C:14:CYS:SG	5:C:201:HEC:HAB	0.67	2.29	7	15
3:C:43:ALA:HB3	3:C:48:TYR:OH	0.66	1.91	6	15
3:C:55:LYS:HD3	3:C:57:VAL:HG13	0.65	1.68	10	15
3:C:87:LYS:O	3:C:90:ASP:HB2	0.65	1.92	7	15
3:C:8:THR:HG22	3:C:11:LYS:HE3	0.64	1.69	2	15
3:C:14:CYS:SG	5:C:201:HEC:HBB3	0.63	2.34	14	15
3:C:22:LYS:NZ	3:C:33:HIS:HD2	0.63	1.92	13	15
3:C:58:LEU:O	3:C:63:ASN:ND2	0.62	2.31	14	15

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
2:B:14:CYS:SG	5:B:201:HEC:C3B	0.62	2.85	14	15
1:A:1:THR:O	1:A:2:THR:HG23	0.61	1.95	14	15
3:C:87:LYS:CB	3:C:90:ASP:OD2	0.61	2.46	15	15
1:A:125:PRO:HG2	1:A:267:GLU:OE1	0.60	1.96	14	15
3:C:17:CYS:SG	5:C:201:HEC:HBC3	0.60	2.36	13	15
3:C:8:THR:CA	3:C:11:LYS:CD	0.60	2.77	8	15
2:B:17:CYS:HG	5:B:201:HEC:CAC	0.60	2.07	14	15
2:B:49:THR:HG22	2:B:52:ASN:H	0.59	1.55	14	15
3:C:1:GLY:O	3:C:96:THR:HG21	0.59	1.98	9	15
3:C:8:THR:CG2	3:C:11:LYS:HE3	0.59	2.28	11	15
2:B:61:GLU:OE2	2:B:99:LYS:NZ	0.58	2.36	14	15
3:C:16:GLN:N	3:C:16:GLN:OE1	0.58	2.35	8	8
3:C:16:GLN:OE1	3:C:16:GLN:N	0.58	2.35	5	7
2:B:70:ASN:ND2	2:B:73:LYS:HG3	0.58	2.14	14	15
1:A:103:SER:OG	1:A:106:ASP:HB2	0.57	1.97	14	15
3:C:91:ARG:O	3:C:95:ILE:HG12	0.56	2.00	14	15
3:C:6:GLY:HA3	3:C:93:ASP:O	0.56	2.01	15	15
1:A:2:THR:CB	1:A:3:PRO:HD2	0.56	2.31	14	15
1:A:292:GLN:O	1:A:294:LEU:HG	0.56	2.00	14	15
1:A:28:LEU:O	1:A:32:GLU:HG3	0.56	2.01	14	15
3:C:5:LYS:O	3:C:7:ALA:N	0.55	2.40	15	15
3:C:8:THR:CB	3:C:11:LYS:HE3	0.55	2.32	13	15
3:C:39:HIS:CD2	3:C:56:ASN:OD1	0.55	2.60	15	15
3:C:79:LYS:HE3	5:C:201:HEC:HMD1	0.54	1.80	15	15
3:C:37:GLY:O	3:C:58:LEU:HD11	0.53	2.04	10	15
1:A:226:LYS:HD2	1:A:227:SER:N	0.53	2.19	14	15
2:B:96:THR:O	2:B:100:LYS:HG3	0.53	2.04	14	15
2:B:61:GLU:HG2	2:B:99:LYS:HD2	0.53	1.79	14	15
1:A:74:LYS:HD2	1:A:74:LYS:C	0.52	2.25	14	9
1:A:74:LYS:C	1:A:74:LYS:HD2	0.52	2.25	8	6
3:C:5:LYS:O	3:C:6:GLY:C	0.52	2.48	10	15
3:C:89:LYS:HB3	3:C:93:ASP:OD2	0.52	2.05	8	15
3:C:55:LYS:HD3	3:C:57:VAL:CG1	0.51	2.36	8	15
3:C:8:THR:O	3:C:11:LYS:HD3	0.51	2.05	15	15
2:B:30:PRO:HG3	2:B:46:TYR:CD2	0.51	2.40	14	15
3:C:22:LYS:C	3:C:24:GLY:H	0.51	2.09	12	15
2:B:17:CYS:SG	5:B:201:HEC:HBC3	0.50	2.46	14	15
3:C:82:PHE:CE2	3:C:84:GLY:HA2	0.50	2.42	10	15
3:C:23:GLY:O	3:C:24:GLY:O	0.50	2.30	13	15
3:C:22:LYS:HD2	3:C:33:HIS:NE2	0.49	2.22	4	15
2:B:49:THR:HG23	2:B:51:ALA:N	0.49	2.22	14	15

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:52:HIS:HE1	1:A:81:SER:O	0.49	1.90	14	15
1:A:226:LYS:HD2	1:A:226:LYS:C	0.49	2.28	14	9
1:A:226:LYS:C	1:A:226:LYS:HD2	0.49	2.28	8	6
2:B:43:ALA:CB	2:B:48:TYR:OH	0.49	2.61	14	15
1:A:288:THR:OG1	1:A:291:GLU:HG3	0.48	2.08	14	15
3:C:42:GLN:O	3:C:43:ALA:C	0.48	2.50	15	15
3:C:55:LYS:O	3:C:56:ASN:C	0.48	2.52	10	15
3:C:7:ALA:HB2	3:C:97:TYR:CD2	0.48	2.44	15	15
2:B:43:ALA:HB3	2:B:48:TYR:OH	0.47	2.09	14	15
1:A:49:LEU:HD13	1:A:49:LEU:C	0.47	2.30	14	10
2:B:7:ALA:O	2:B:11:LYS:HG3	0.47	2.09	14	15
3:C:18:HIS:CE1	3:C:30:PRO:HD2	0.47	2.45	12	15
1:A:49:LEU:C	1:A:49:LEU:HD13	0.47	2.30	8	5
3:C:43:ALA:CB	3:C:48:TYR:OH	0.47	2.62	10	15
2:B:53:ILE:C	2:B:53:ILE:HD13	0.47	2.30	14	8
2:B:53:ILE:HD13	2:B:53:ILE:C	0.47	2.30	10	7
1:A:93:GLU:OE1	1:A:94:PRO:HD3	0.46	2.10	14	15
2:B:50:ASP:O	2:B:53:ILE:HG22	0.46	2.11	14	15
3:C:15:LEU:O	3:C:16:GLN:C	0.46	2.54	13	15
1:A:58:ASP:HA	1:A:160:ARG:CD	0.46	2.38	14	15
3:C:14:CYS:SG	5:C:201:HEC:C3B	0.46	3.02	14	15
3:C:8:THR:HB	3:C:11:LYS:HE3	0.46	1.87	11	15
2:B:102:THR:O	2:B:103:GLU:OXT	0.45	2.34	14	15
2:B:30:PRO:HG3	2:B:46:TYR:CE2	0.45	2.46	14	15
3:C:23:GLY:O	3:C:24:GLY:C	0.45	2.55	5	15
3:C:55:LYS:O	3:C:56:ASN:O	0.45	2.34	15	15
3:C:18:HIS:CD2	5:C:201:HEC:NB	0.45	2.85	13	15
1:A:2:THR:HG22	1:A:3:PRO:HD2	0.45	1.87	14	15
1:A:35:GLU:HG2	1:A:39:TYR:CZ	0.45	2.46	14	15
3:C:-2:LYS:O	3:C:-1:ALA:HB2	0.45	2.12	12	15
2:B:17:CYS:SG	5:B:201:HEC:HAC	0.44	2.43	14	15
2:B:70:ASN:CG	2:B:73:LYS:HG3	0.44	2.33	14	15
1:A:2:THR:CG2	1:A:3:PRO:HD2	0.43	2.44	14	15
1:A:134:PRO:HG2	1:A:136:ASP:OD1	0.43	2.14	14	15
3:C:22:LYS:HZ3	3:C:33:HIS:HD2	0.43	1.56	14	5
3:C:98:LEU:O	3:C:99:LYS:C	0.43	2.57	6	15
3:C:89:LYS:O	3:C:93:ASP:OD2	0.42	2.37	8	15
3:C:5:LYS:O	3:C:8:THR:N	0.42	2.52	14	15
1:A:52:HIS:CE1	1:A:81:SER:O	0.42	2.72	14	15
3:C:77:GLY:O	3:C:78:THR:C	0.42	2.58	13	11
2:B:68:LEU:HD13	2:B:94:LEU:HD23	0.41	1.90	14	15

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
3:C:26:HIS:NE2	3:C:44:GLU:O	0.41	2.53	13	11
3:C:22:LYS:NZ	3:C:33:HIS:CD2	0.41	2.82	14	14
3:C:97:TYR:C	3:C:97:TYR:CD1	0.41	2.94	7	13
3:C:60:ASP:O	3:C:62:ASN:N	0.41	2.54	13	13
3:C:97:TYR:CD1	3:C:97:TYR:C	0.41	2.94	13	2
2:B:88:GLU:OE1	2:B:91:ARG:NH1	0.41	2.53	14	15
3:C:33:HIS:CE1	3:C:103:GLU:OXT	0.41	2.74	14	15
3:C:89:LYS:O	3:C:93:ASP:N	0.41	2.44	14	2
2:B:41:GLY:HA2	2:B:48:TYR:CD2	0.40	2.52	14	15
3:C:35:ILE:HA	3:C:35:ILE:HD12	0.40	1.79	15	1
3:C:89:LYS:C	3:C:93:ASP:OD2	0.40	2.60	15	1

6.3 Torsion angles (i)

6.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	292/294 (99%)	285±0 (98±0%)	5±0 (2±0%)	2±0 (1±0%)	26	73
2	B	106/108 (98%)	100±0 (94±0%)	6±0 (6±0%)	0±0 (0±0%)	100	100
3	C	106/108 (98%)	91±0 (86±0%)	9±0 (8±0%)	6±0 (6±0%)	3	22
All	All	7560/7650 (99%)	7140 (94%)	300 (4%)	120 (2%)	13	57

All 8 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
3	C	6	GLY	15
1	A	13	GLY	15
3	C	56	ASN	15
3	C	24	GLY	15
3	C	-1	ALA	15
1	A	3	PRO	15
3	C	61	GLU	15
3	C	-2	LYS	15

6.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	252/252 (100%)	245±0 (97±0%)	7±0 (3±0%)	46	90
2	B	89/89 (100%)	86±0 (97±0%)	3±0 (3±0%)	40	87
3	C	89/89 (100%)	73±0 (82±0%)	16±0 (18±0%)	4	38
All	All	6450/6450 (100%)	6060 (94%)	390 (6%)	23	72

All 26 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
3	C	98	LEU	15
3	C	55	LYS	15
1	A	166	ARG	15
3	C	61	GLU	15
3	C	-2	LYS	15
3	C	99	LYS	15
3	C	80	MET	15
3	C	-4	GLU	15
3	C	88	GLU	15
1	A	2	THR	15
3	C	2	SER	15
1	A	3	PRO	15
1	A	215	LYS	15
3	C	22	LYS	15
3	C	-3	PHE	15
3	C	-5	THR	15
2	B	50	ASP	15
3	C	11	LYS	15
1	A	4	LEU	15
1	A	226	LYS	15
1	A	74	LYS	15
3	C	58	LEU	15
2	B	53	ILE	15
2	B	49	THR	15
3	C	85	LEU	15
3	C	73	LYS	15

6.3.3   RNA   (i)

There are no RNA molecules in this entry.

6.4   Non-standard residues in protein, DNA, RNA chains   (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.5   Carbohydrates   (i)

There are no carbohydrates in this entry.

6.6   Ligand geometry   (i)

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds for which Mogul statistics could be retrieved, the number of bonds that are observed in the model and the number of bonds that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length is the number of standard deviations the observed value is removed from the expected value. A bond length with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond lengths.

Mol	Type	Chain	Res	Link	Bond lengths		
					Counts	RMSZ	#Z>2
5	HEC	B	201	2	26,50,50	1.47±0.00	0±0 (0±0%)
5	HEC	C	201	3	26,50,50	2.66±0.00	2±0 (7±0%)
4	HEM	A	301	1	27,50,50	1.94±0.00	0±0 (0±0%)

In the following table, the Counts columns list the number of angles for which Mogul statistics could be retrieved, the number of angles that are observed in the model and the number of angles that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond angle is the number of standard deviations the observed value is removed from the expected value. A bond angle with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the average root-mean-square of all Z scores of the bond angles.

Mol	Type	Chain	Res	Link	Bond angles		
					Counts	RMSZ	#Z>2
5	HEC	B	201	2	18,82,82	1.74±0.00	0±0 (0±0%)
5	HEC	C	201	3	18,82,82	2.49±0.01	0±0 (0±0%)
4	HEM	A	301	1	17,82,82	2.09±0.00	1±0 (5±0%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	HEC	C	201	3	-	0±0,6,54,54	-
5	HEC	B	201	2	-	0±0,6,54,54	-
4	HEM	A	301	1	-	0±0,6,54,54	-

All unique bond outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)	Models	
								Worst	Total
5	C	201	HEC	C3C-C2C	8.30	1.32	1.40	12	15
5	C	201	HEC	C3B-C2B	6.18	1.34	1.40	15	15

All unique angle outliers are listed below.

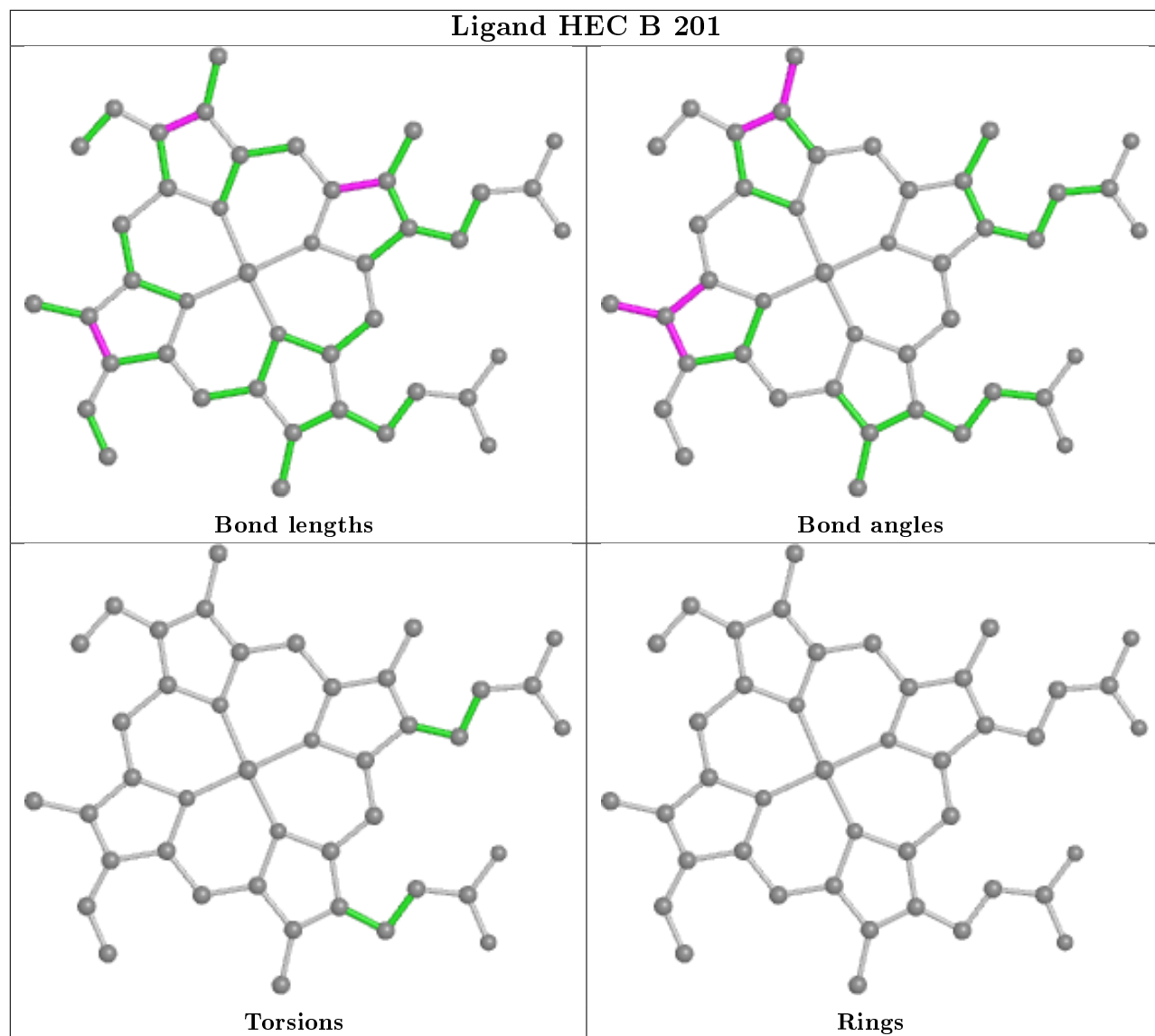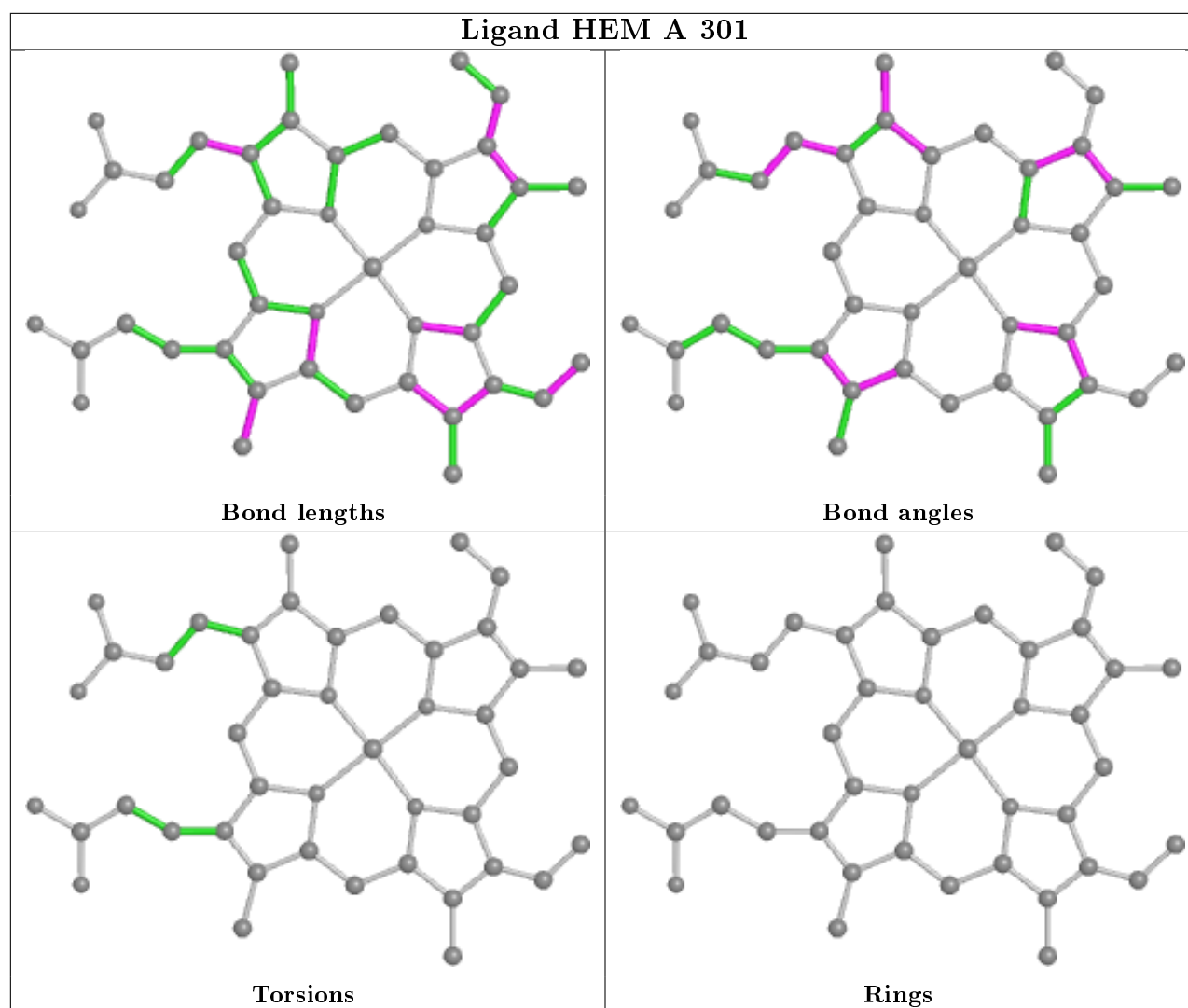| Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
4	A	301	HEM	C4A-C3A-C2A	5.34	103.28	107.00	11	15

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand HEC C 201

Bond lengths

Bond angles

Torsions

Rings

Ligand HEC B 201

Bond lengths

Bond angles

Torsions

Rings

Ligand HEM A 301

Bond lengths

Bond angles

Torsions

Rings
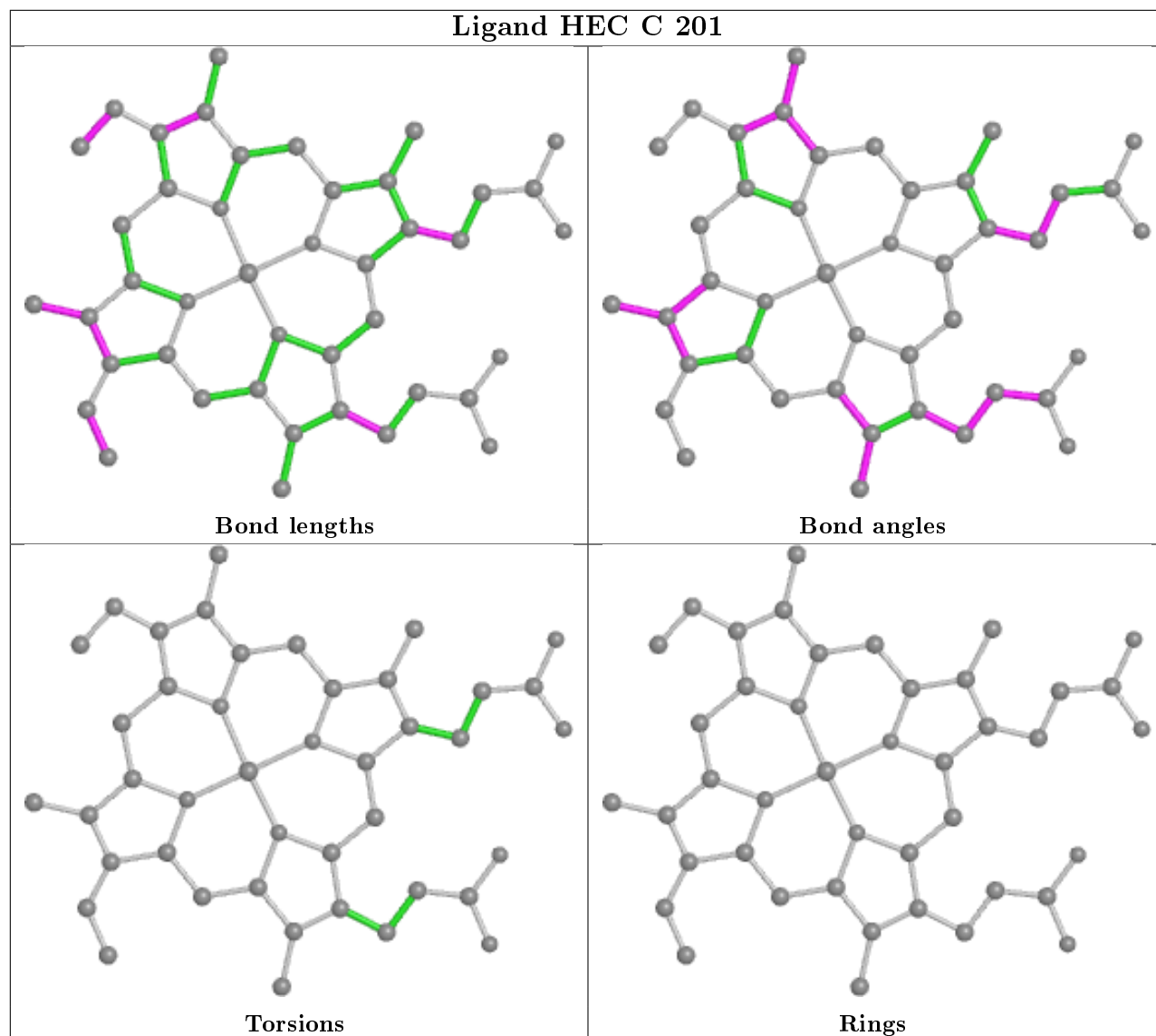
6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation (i)

The completeness of assignment taking into account all chemical shift lists is 7% for the well-defined parts and 7% for the entire structure.

7.1 Chemical shift list 1

File name: input_cs.cif

Chemical shift list name: *assigned_ chem_ shift_ list_ 1*

7.1.1 Bookkeeping (i)

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	458
Number of shifts mapped to atoms	458
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.1.2 Chemical shift referencing (i)

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	0	—	None (insufficient data)
$^{13}C_{\beta}$	0	—	None (insufficient data)
$^{13}C'$	0	—	None (insufficient data)
^{15}N	229	0.35 ± 0.29	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments (i)

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 7%, i.e. 458 atoms were assigned a chemical shift out of a possible 6311. 0 out of 60 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	458/2502 (18%)	229/996 (23%)	0/1020 (0%)	229/486 (47%)
Sidechain	0/3166 (0%)	0/1870 (0%)	0/1146 (0%)	0/150 (0%)

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	0/643 (0%)	0/336 (0%)	0/273 (0%)	0/34 (0%)
Overall	458/6311 (7%)	229/3202 (7%)	0/2439 (0%)	229/670 (34%)

The following table shows the completeness of the chemical shift assignments for the full structure. The overall completeness is 7%, i.e. 458 atoms were assigned a chemical shift out of a possible 6311. 0 out of 60 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	458/2502 (18%)	229/996 (23%)	0/1020 (0%)	229/486 (47%)
Sidechain	0/3166 (0%)	0/1870 (0%)	0/1146 (0%)	0/150 (0%)
Aromatic	0/643 (0%)	0/336 (0%)	0/273 (0%)	0/34 (0%)
Overall	458/6311 (7%)	229/3202 (7%)	0/2439 (0%)	229/670 (34%)

7.1.4 Statistically unusual chemical shifts ⓘ

There are no statistically unusual chemical shifts.

7.1.5 Random Coil Index (RCI) plots ⓘ

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.

Random coil index (RCI) for chain A: